"NOVEL HYBRID MACHINE LEARNING ALGORITHM DEVELOPMENT FOR CARDIAC TREATMENT WITH RECOMMENDATION SYSTEM "

SYNOPSIS

Submitted By

MARIUM SHUJAUDDIN SHAIKH

For The Degree Of

DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE

UNDER THE GUIDANCE
OF

DR.H.S. FADEWAR
ASSISTANT PROFESOR
SCHOOL OF COMPUTATIONAL SCIENCES
SWAMI RAMANAND TEERTH MARATHWADA UNIVERSITY, NANDED

SWAMI RAMANAND TEERTH MARATHWADA UNIVERSITY, NANDED

TABLE OF CONTENT

Chapter	Particular	Page No
1	ABSTRACT	3
2	INRTODUCTION	4
3	LITERATURE REVIEW	5
4	STATEMENT OF PROBLEM	7
5	OBJECTIVES OF RESEARCH	8
6	METHODOLOGY TO BE ADOPTED	8
7	IMPORTANCE OF STUDY	9
8	PROPOSED WORK PLAN	10
9	REFERENCES	9

“NOVEL HYBRID MACHINE LEARNING ALGORITHM DEVELOPMENT FOR CARDIAC TREATMENT WITH RECOMMENDATION SYSTEM “

1.Abstract

Heart Attack, Heart failure and other Heart disease are very common reason of death in India. Ischemic heart disease and stroke are the predominant causes and are responsible for >80% of CVD deaths according to WHO. After covid-19 heart attack rate are increasing even in middle age of group of people. Heart disease become one of the leading reasons of death. History of patient disease , life style of person and stress level are the same factors which contribute to heart disease. Most of Medical Data are available but not in organized and structured form which cause delay in detection and timely treatment. Machine Learning will help to link up those scattered data and Help Medical experts to provide proper diagnosis on time. Purpose of this study to optimize prediction of heart disease using ML.

Here we are considering 3 sets of classification algorithms XGboost, Regression Analysis and SVM . which will be best for modeling for heart disease detection analysis will be based on mainly ECG and some other important feature.

This prediction can help clinically in analyzing the risk factors of the disease and interpretation of the patient scenario

Keywords: CVD(Cardiovascular disease),ECG(Electro Cardio diagram),Machine Learning ,XGBoost, Regression Analysis ,Classification Algorithms.

2-INTRODUCTION:

An electrocardiogram (ECG) is a simple test that can be used to check your heart's rhythm and electrical activity **,** Sensors attached to the skin are used to detect the electrical signals produced by your heart each time it beats.

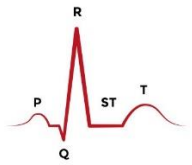
An ECG can help detect arrhythmias ,coronary heart disease, heart attacks and cardiomyopathy

There are 3 main types of ECG:

A resting ECG – carried out while you're lying down in a comfortable position

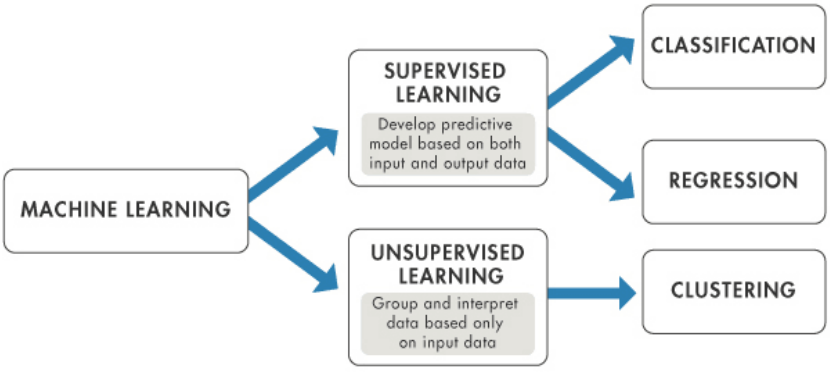
A stress or exercise ECG – carried out while you're using an exercise bike or treadmill

an ambulatory ECG (sometimes called a Holter monitor) – the electrodes are connected to a small portable machine worn at your waist so your heart can be monitored at home for 1 or more days



- **P Wave**: Initial wave
- **QRS Complex**: As the name suggests, the QRS complex includes the Q wave, R wave, and S wave. These three waves occur in rapid succession.
- **T Wave**: A T wave follows the QRS complex and indicates ventricular repolarization. Unlike a P wave, a normal T wave is slightly asymmetric

this research which help to build system which can detect abnormalities in heartbeat on timely manners and prevent sudden heart attacks using classification based supervised machine learning algorithm such as XGBoost, Logistic Regression and Support vector Machine. we will train Machine by feeding different reading and classified the heart beat abnormality based on ECG data. Supervised machine learning learns patterns and relationships between input and output data. It is defined by its use of labelled data. A labelled data is a dataset that contains a lot of examples of Features and Target. Supervised learning uses algorithms that learn the relationship of Features and Target from the dataset. This process is referred to as Training or Fitting.



Classification is a type of supervised machine learning where algorithms learn from the data to predict an outcome or event in the future. Where features and target will be used achieved objectives. Features will be attributes of a Sensitivity, Specificity, Precision etc. and Target will represent output normally represented by 1 or 0 / True or False / Yes or No.

3-Literature review:

A thorough search has been done of the previous work on the domain of the heart disease using different algorithms**.** The inclusion criteria was based on identifying the papers which are of related domain, utilization of latest machine learning algorithms, challenging area in domain of heart disease. Search terms for identifying a “Challenges in identifying health disease” “Heart disease using Machine Learning” and “comparison

between different classification algorithms". The exclusion criteria included removing duplicate papers, papers which presented inferior work in terms of evaluation parameter values, and obsolete work.

Different types of Machines Learning oriented research papers have been analyzed According to Nadal N,Goel L and Tanwar R More featured medical attributes can be used for providing better model with more accuracy and performance. A good understanding of the required parameters for the diagnosis of the disease can be highly helpful in making precise and accurate predictions.- Cardiovascular (CV) disease research and treatment coupled with some high-performance tools for analysis can improve the knowledge about the domain.

According to Hetal Bhavsar, Amit Ganatra The model implemented in SVMs is determined by the choice of the kernel. Deciding which kernel is the most suitable for a given application is obviously an important (and open) issue. Expansion to very large database includes a large proportion of the training data, which leads to a model that is expensive both to store and to evaluate. Alleviating this problem is one area of ongoing research in SVMs.

Ramalingam, V. V., Ayantan Dandapath, and M. Karthik Raja has conducted survey where they found SVM,K- K-Nearest Neighbour (KNN), NaïveBayes, Decision Trees (DT), Random Forest (RF) and ensemble models are found very popular among the researchers.

Medhekar, Dhanashree S., Mayur P. Bote, and Shruti D. Deshmukh says Accuracy of system depends upon algorithm and database used.

In study Bhatt CM, Patel P, Ghetia T, Mazzeo PL says it would be beneficial to evaluate the performance of the model on a held-out test dataset in order to establish its generalizability to new, unseen data. Ultimately, future research should aim to establish the robustness and generalizability of the results and the interpretability of the clusters formed by the algorithm, which could aid in understanding the results and support decision

making based on the study's findings.it would be beneficial to evaluate the performance of the model on a held-out test dataset in order to establish its generalizability to new, unseen data. Ultimately, future research should aim to establish the robustness and generalizability of the results and the interpretability of the clusters formed by the algorithm, which could aid in understanding the results and support decision making based on the study's findings.

Umarani Nagavelli, Debara Samanta, Pratha Chakraborty suggest that the accuracy parameter is high in the XGBoost algorithm-based heart disease detection which can be updated in the future by adding more attributes to the heart disease dataset and making it more interactive for the users. It can also be carried out as a mobile application with reduced computing time and complexity.

Subramani S, Varshney N, Anand MV, Soudagar MEM, Al-keridis LA, Upadhyay TK, Alshammari N, Saeed M, Subramanian K, Anbarasu K and Rohini K (2023) finding suggests that stacking models can combine the benefits of a variety of model types to achieve superior prediction performance The general conclusion drawn from the implemented models are listed below:

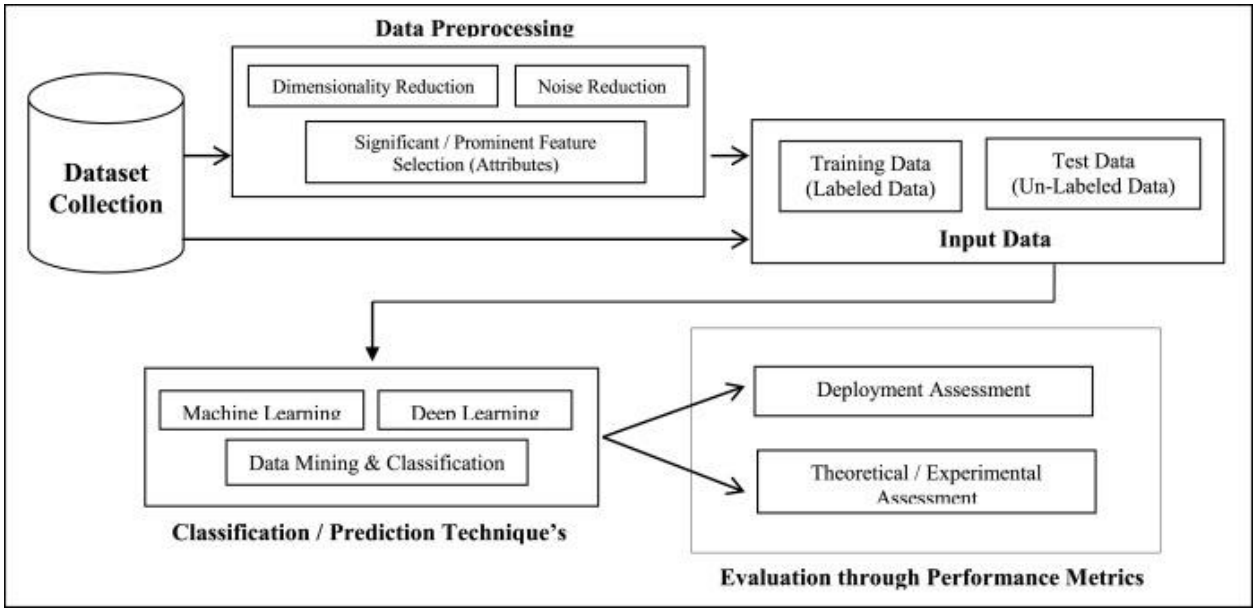
(1) Noise in sensors image can produce not accurate reading. Filter is needed while reading P,Q,R,S,T values from ECG graph.

(2) Continuous information can be used in place of categorical information to build a heart disease system with early detection.

4-Problem of the Statement:

The Main objective of the work analyze which algorithm model will work best for heart disease detection based on ECG data. Through Data we can extract different features to analyze and

compares the models so we can understand which model can give better diagnosis and prediction of heart deceased and its complications.



5-Research Objectives:

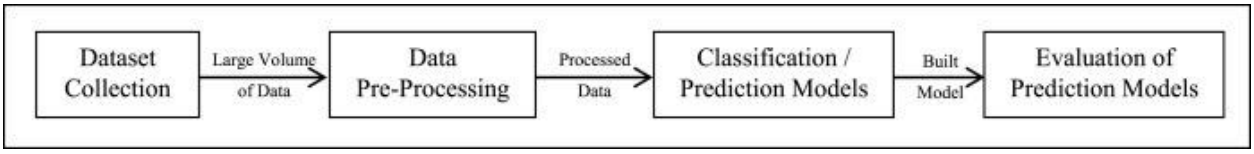
1. Study and compare different Machine Learning training model algorithms which are suitable to detect CVD.
2. Develop hybrid algorithm based on analysis.

6-Research Methodology

Experimental research design: This research design is used to study causal relationships. One or more independent variables are manipulated, and their effect on one or more dependent variables is measured.

Correlational research design: Correlation reflects the strength and/or direction of the relationship among variables. The direction of a correlation can be positive or negative. Correlational research

design helps researchers establish a relationship between two variables without the researcher controlling any of them.



Medical data are available in huge form but very unstructured. Data can be collected from online resource Physionet , keggel also in talk with one hospital so it can be performed on live data. If Accuracy and Efficiency among models can be determined it can help to choose to best model which can give accurate result.

Tools for prediction/classification models

With the help of python framework/MATLAB /Google colab we can clean data, extract features and execute models and compare different features based on objectives.

Various studies are reporting that the Mahine learning will improve cardiovascular disease detection. The publicly available datasets of heart disease such as Clevel and, keggel, physionet and Statlog are mostly used for comparing the prediction models' performance between researchers

7- **Importance of study:**

CVD is leading cause of death, ML will help to make data in more structured format, makes use of unused data and proper data analysis with machine reduce human errors and Help medical experts to detect and diagnose heart issues with more accurate and precise manner

8-**Propsed Work Plan**

1-Selection of 3 Proper Algorithms which are suitable for CVD.

2-Collection Of data

3-Filter Data

4-Extract Features

5-Test Data

6-Train Machine With the help data

7-Evaluate Result

9-REFERENCES

[1]. Nandal N, Goel L and TANWAR R. Machine learning-based heart attack prediction: A symptomatic heart attack prediction method and exploratory analysis [version 1; peer review: awaiting peer review]. *F1000Research* 2022, 11:1126

[2]Batta Mahesh ,Machine Learning Algorithm A review. International Journal of Science and Research (IJSR) ISSN: 2319-7064 ResearchGate Impact Factor (2018)

[3]Hetal Bhavsar, Amit Ganatra. A Comparative Study of Training Algorithms for Supervised Machine Learning International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-4, September 2012

[4] Ratnasari NR, Susanto A, Soesanti I, *et al.*: Thoracic X-ray features extraction using thresholding-based ROI template and PCA-based features selection for lung TB classification purposes. *Proceedings of the 2013 3rd International Conference on Instrumentation, Communications, Information Technology and Biomedical Engineering (ICICI-BME); Bandung, Indonesia. IEEE* November 2013; pp. 65–69.

[5] Dorairaj Prabhakaran, Panniyammakal Jeemon and Ambuj Roy Cardiovascular Diseases in IndiaCurrent Epidemiology and Future Directions riginally published19 Apr 2016

[6]Negi S, Kumar Y, Mishra VM: Feature extraction and classification for EMG signals using linear discriminant analysis. Proceedings of the 2016 2nd International Conference on Advances in Computing, Communication, & Automation (ICACCA) (Fall); September 2016; Bareilly, India. IEEE.

[7] Ramalingam, V. V., Ayantan Dandapath, and M. Karthik Raja. "Heart disease prediction using machine learning techniques: a survey." *International Journal of Engineering & Technology* 7.2.8 (2018): 684-687.

[8] Medhekar, Dhanashree S., Mayur P. Bote, and Shruti D. Deshmukh. "Heart disease prediction system using naive Bayes." *Int. J. Enhanced Res. Sci. Technol. Eng* 2.3 (2013).

[9] Bhatt CM, Patel P, Ghetia T, Mazzeo PL. Effective Heart Disease Prediction Using Machine Learning Techniques. *Algorithms*. 2023; 16(2):88. https://doi.org/10.3390/a16020088

[10] Nagavelli, Umarani et al. "Machine Learning Technology-Based Heart Disease Detection Models." *Journal of healthcare engineering* vol. 2022 7351061. 27 Feb. 2022, doi:10.1155/2022/7351061

[11]Subramani S, Varshney N, Anand MV, Soudagar MEM, Al-keridis LA, Upadhyay TK, Alshammari N, Saeed M, Subramanian K, Anbarasu K and Rohini K (2023) Cardiovascular diseases prediction by machine learning incorporation with deep learning. *Front. Med*. 10:1150933. doi: 10.3389/fmed.2023.1150933

Research Scholar Guide